JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1971, 16, 283-288 NUMBER 2 (SEPTEMBER)

# *AVERSIVE CONTROL: A REVIEW OF B. A. CAMPBELL AND R. M. CHURCH'S* PUNISHMENT AND AVERSIVE BEHAVIOR[1]

ROBERT C. BOLLES

UNIVERSITY OF WASHINGTON

This book is the product of the conference on punishment held at Princeton in the summer of 1967. It is a much more successful book than are most published conference meetings. Some of the reasons for this success are that the hosts picked the participants carefully, they set the scope of the conference appropriately, and as editors they have done a great deal of work in getting the manuscripts into shape for publication. Although some of the individual papers are more impressive than others, all are well written and of high quality. The different papers are relatively homogeneous in terms of level, length, abstruseness, and they all present a good balance of data and discussion. Finally, Campbell and Church have done a useful service in gathering the transcript of the discussion that occurred during the conference and boiling it down to a few pages of appendix at the end of the book. This appendix constitutes, in effect, a separate chapter that is of considerable value in summarizing some important methodological and conceptual points. Usually, conference discussions are given more or less verbatim and the points get broken up into pieces, scattered around, and watered down. A fairly complete bibliography on punishment, compiled and keyed by Boe, has been tacked on with pomp and ceremony at the end.

The book is successful also in that it provides some sense of closure on many of the topics covered. Such closure also contrasts with the typical appearance of published conference proceedings in which the author appears bent upon awing his audience with his latest tentative findings. The impression typically gained is that an immense task confronts workers in the subject and anyone else who wants to understand. Here, although punishment is a broad area, and the subject matter of the conference is still broader, we have the impression that a number of men have obtained a pretty good understanding of their problems. If many of the problems are not yet resolved, there is some sense that at least we are making progress on them, that the issues are fairly clear, and that many pieces of the puzzle are beginning to fall into place. There is, in short, an unusual and refreshing sense of optimism about the book. This is not to say that we are on the brink of solving all the problems presented by the punishment procedure. My view of it is, rather, that there is a certain finite number of problems and that we have some chance of coming to grips with them without being overwhelmed by the complexity of the subject matter. Perhaps the area of punishment is becoming increasingly subject to the laws of behavior, in contrast with many areas of psychology that seem to be becoming increasingly complex as more investigators discover more and more curious phenomena.

This optimistic tone is set in the first chapter in which Campbell and Masterson describe a series of studies that extend the pioneering work of Campbell and Teghtsoonian (1958). The procedure involves a standardized tilt cage in which shock of one type and intensity is presented when the animal is on one side, and a different shock is presented when the animal is on the other side. Under most pairs of conditions, rats soon stabilize the proportion of time they spend on a given side. This choice tells us something about the sensory dimension of shock as well as something about its aversiveness.

In Campbell and Teghtsoonian's original study, it looked as though different kinds of shock had qualitatively as well as quantitatively different effects upon behavior. Differ-

[1]New York: Appleton-Century-Crofts, 1969. Pp 597. $14.45.

ent kinds of shock generators led to different kinds of adjustive behavior, and to different rat resistances, and all the parameters changed over time. It appeared that we were all cursed if we did not use the appropriate kind of shock for a given kind of behavioral control. But here, primarily because of better data analysis, the results with the tilt cage led to much tidier and simpler conclusions. A reasonably good first approximation to the data is that what matters is the current flowing through the animal. As long as one does not use too high a voltage or low a series resistor, then it does not seem to make much difference whether one uses a constant current source or a fixed impedance source. The current flowing through the animal seems to determine the aversiveness of the shock, its discriminability, and its effectiveness as an aversive stimulus.

Some additional small good things also come out of these data. One is that the behavioral data are nearly the same after a minute or two as after hours of testing. A second is that the rat shows no learning to deal with shock; at least it does not increase its resistance over time. Another is that the aversion threshold is very near the detection threshold. There is no evidence in any of Campbell's work that shock is ever a positive stimulus; as soon as the rat detects it, it shows an aversion to it.

Another positively toned chapter is the one by Hoffman, which is concerned primarily with the establishment of discriminative control over the CER. Hoffman starts by examining in considerable detail the phenomenon of conditioned suppression. He describes a variety of functional relationships with different experimental parameters so that the uninitiated reader has a good sense of what the phenomenon looks like. This is followed up by a detailed analysis of discriminative condtioning using a variety of techniques, especially fade-in. This chapter, together with Church's chapter, provide a thorough introduction to the suppression literature.

Church's chapter is supposed to be an introduction to the general questions on conditioned suppression. His data (mostly from his own laboratory) are impressive, his logical analysis is cool and sound, and he arrives at some general conclusions, which include the following: (1) response suppression increases with the severity of shock, both intensity and duration, (2) response suppression increases with the temporal contiguity of the punished response and shock, (3) response suppression is greater if there is a contingency between the response and shock, *i.e.*, other things being equal, contingent shock is more effective than non-contingent shock in suppressing a response, (4) suppression of a response sequence is greater if an early member of a chain is punished than if a later response is punished. These conclusions are stated with considerable assurance, but for various reasons the whole account lacks conviction. Part of my uneasiness is probably due to the fact that Church's arguments are not derived from the entire suppression literature but mainly from studies done in his laboratory using relatively standardized procedures. What are the boundary conditions within which they will continue to summarize research findings? There is some irony in the contrast between the assurance with which Church tells us here that contingent shock is more suppressive than non-contingent shock and the doubt raised by the recent report from his laboratory (Church, Wooten, and Matthews, 1970) which appears to throw the question wide open again. Secondly, I am uncomfortable with Church's a-theoretical style of analysis. Perhaps there are readers who can follow with interest his arguments based upon purely empirical and methodological considerations, but I cannot. I seem to need reference to the underlying psychological processes. For example, why is it true—where it is true—that response-contingent shock produces greater suppression?

These three chapters survey their own domain rather extensively, and also provide a sort of foundation for the remaining topics. The editors have organized the chapters into groups, of course, but the groupings themselves do not account for much of the variance among the papers, so I will refer to them in terms of my own arbitrary groupings. There is a group concerned with avoidance learning, another bearing on the theory of punishment and/or its applications to the human condition, and a final group dealing with what I would call "special topics".

One chapter that does not fall into this classification is the one by Rachlin and Herrnstein. The high point of their paper is the title, "Hedonism Revisited". Their animals were given a choice period in which respond-

ing on one key led to a program that provided fixed-interval food plus some shock schedule, while responding on the other key led to a program with the same food payoff but a different shock schedule. By systematically varying the shock contingencies on the two programs it should be possible to measure their relative aversiveness in terms of choice behavior in the two-key choice situation. The functions come out surprisingly flat, however, and surprisingly insensitive to changes in the shock schedules. The reason for this may be that in the choice situation the transitions to the alternative programs were themselves on VI 3-min schedules. This feature of the study was introduced so that even when there was a low rate on a given key there would be a good chance that the program would be initiated. But doesn't this procedure, in effect, deny the birds a choice? The major disappointment in the chapter is that the authors do not make clear what their birds' behavior tells us about hedonism. Hedonism has always had reference to a particular kind of hypothetical process, usually with mentalistic overtones. On the other hand, Rachlin and Herrnstein are among those concerned with the effective control of behavior, rather than with finding hypothetical explanatory processes. Thus, we might expect them to assert that hedonistic principles can be used heuristically to predict how animals will respond to differential outcomes. But, as they stand now, hedonistic principles are surely too general, rough, and heuristic to be useful either to those who seek effective ways of controlling behavior or to those who search for explanatory mechanisms.

Let us turn now to the several chapters that deal with "special topics". A chapter by Wagner is concerned with whether withdrawal of positive reinforcement is aversive. There is a growing tendency for operant workers to assume that it is, and to consider loss of reinforcement, reduction in reinforcement, or even timeout as aversive events. The results with multiple and chain schedules indicate that, at least procedurally, frustration is aversive. However, as Leitenberg (1965) has ably argued, the introduction of a possibly aversive or frustrative factor is confounded by a reduction in positive reinforcement. We really cannot tell whether the animal is avoiding withdrawal of food or simply maximizing its food intake. Thus, the problem is of considerable methodological and conceptual importance. Wagner discusses the issues at some length and notes that most of the experiments done in the discrete trial tradition have failed to show that frustration is aversive. In a study by Brown and Wagner, resistance to extinction was found to be enhanced both for groups trained with nonreward or with punishment (shock) and for groups switched in extinction from one to the other. In each case extinction performance was comparable, which, Wagner suggests, means that the aversiveness of nonreward and punishment are comparable. This tortuous kind of argument was all there was until Daly (1969) demonstrated rather conclusively that rats will learn a new response to escape from situational cues that had been paired with withdrawal of reinforcement. Curiously, the question today seems to be not so much whether withdrawal of reinforcement is aversive as whether it is motivating (Staddon, 1970). This area has changed a good deal since Wagner's chapter was written, so that it becomes primarily a historical introduction to the problem.

One of the best written and most exciting chapters is that by Hearst on the factors involved in the stimulus control of suppression. Hearst starts with an interesting paradox: in spite of what the textbooks say about the sharp generalization gradient of avoidance and the flat gradient for approach behavior (results generally found in the runway), Hearst found just the opposite. He found a flat gradient for the avoidance response and a sharp gradient for the positively reinforced response in a concurrent appetitive-avoidance operant task. Over the years Hearst has analyzed the situation experimentally, attacking one after another the different procedural and methodological factors involved. His conclusion is that there is nothing intrinsic in appetitive or aversive tasks that leads to flatter or steeper generalization gradients. Rather, the shape of the gradient seems to be a function of the salience of the cues, the number of cues involved, and the amount of training on different cues. It all hangs together, and we are left with the impression that there are basic associative principles in psychology that transcend particular tasks, particular kinds of motivation, and particular response measures. This is one of the fundamental insights that

appear here and there throughout the book. Punishment and aversive procedures in general do not constitute a special realm all their own; they do not require new laws of behavior for their explanation. To a large extent, aversive stimuli are simply stimuli with special stimulus properties. Once the purely associative aspects of the shock, or loud noise, or whatever, are understood, the mystery and peculiarity of the aversive situation disappears.

This point is brought into sharp focus in the beautiful chapter of Fowler and Wischner. They addressed themselves to another paradoxical effect of punishment, namely, that under some conditions animals will learn a discrimination better if mild punishment is made contingent upon the correct response. This is an old paradox, and the literature is complex and rather messy. But Fowler and Wischner lead us through it step by step; they identify the relevant variables, and they test them. They reject the original hypothesis that shock slows the animal down so that it has a better chance to "attend" to the relevant discriminada. In its place emerges the sublime idea that shock has characteristics other than aversiveness. One of shock's characteristics is that it is attention-getting. Then, by being associated with the highly salient shock stimulus, the discriminative stimuli for the correct response become more salient themselves and easier for the animal to respond to. The phenomenon seems to be similar to mediation by verbal labelling in the human subject. Thus, the paradoxical facilitation of discrimination learning looks like a paradox when we think of shock only as aversive—just a thing to be avoided. If we regard shock as a stimulus with a multitude of properties, then the facilitation occurs when we experimentally pit the salience of shock, *i.e.*, its associative value, against its aversiveness in such a way that the former effect predominates.

A similar interpretation could be made of the paradoxical vicious-circle phenomenon. This is the effect that a running response can be maintained, and may actually be strengthened, in extinction if shock is presented in a portion of the runway. It could be said that what this procedure does is pit the reinforcing value of shock termination against the aversiveness, the to-be-avoidedness, of the shock. Under a fairly wide range of experimental conditions, the reinforcing value of shock termination appears to be stronger than the tendency to avoid the shock and the animal is then paradoxically emitting a response that is unnecessary, unwarranted, and actually punished. Brown discusses the phenomenon at great length, but he does not seem to see anything very systematic or lawful about it. It is in a class by itself; self-punitive behavior is like masochism in the neurotic; the laws that apply to it are to be found by parametric studies, which show repeatedly how paradoxical the paradox is. Brown makes no attempt to relate the phenomenon to what we know about shock, or punishment, or extinction, or fear conditioning, or anything else. This is a disappointing chapter.

One more special topic was discovered some years ago by Kamin. One stimulus, $S_1$, is used as a CS in a conditioned suppression situation. A few such trials are then followed by trials with the compound stimulus $S_1 + S_2$ in the same situation. Then the suppressive effects of $S_2$ alone are tested. What Kamin finds is that the initial $S_1$ trials block the subsequent conditioning of fear to $S_2$. At first, Kamin attributed the blocking to an attentional process: attention is initially focused on $S_1$, and $S_2$ is ignored. But here he describes a series of studies that have led to a better and more profound hypothesis, namely, that $S_2$ was redundant in the stimulus compound; it did not predict anything that $S_1$ did not already predict. The block can be broken in a number of ways such as making the compound stimulus predict a more intense shock. Evidently any procedural alteration that gives $S_2$ predictive value makes it an effective stimulus as subsequently measured in a test for suppression.

Thus, Kamin reminds us of a fundamental point regarding the nature of punishment, a point which has been around for some time but which keeps disappearing from view. The point is simply that what counts in the suppression of behavior is the predictability of the primary aversive stimulus. Shock itself will always suppress most ongoing behavior if for no other reason than that it elicits so powerfully its own unconditioned reactions. What all the experimentation is about is to clarify the conditions under which stimuli associated with shock will also suppress ongoing behavior. For me, the reference paradigm is the experiment published by Azrin in

1956 in which he showed that suppression of bar pressing for food was correlated with whatever temporal and exteroceptive stimuli were available to the animal to predict shock. If shock was scheduled for a given period and not at other times, then bar pressing was suppressed in the given period and not at other times. To demonstrate this fundamental principle it is only necessary to give an animal sufficient training that the predictive events (including its own behavior) are discriminated. Given sufficient training, ongoing behavior will be suppressed in the presence of whatever events predict shock, and not suppressed in the presence of other events. Kamin does not expand here upon the importance of the predictability principle, nor do the other conference participants discuss it explicitly. But it was evidently "in the air" at the time. For example, Seligman (1968) has subsequently published a paper that illustrates the principle quite nicely.

Having slipped over from "special topics" to theory, I should note that there are two provocative chapters devoted explicitly to theory. Logan begins his chapter by suggesting that punishment is simply a negative incentive. This is only part of the story, however, because for Logan, incentive motivation does not just modulate behavior, it is the all-important determinant of behavior strength. There is, of course, a complementarity between reinforcement and motivation, such that if a theorist is willing to place enough stress upon the one set of concepts and principles, then he has no need for the other. Logan has placed sufficient emphasis upon incentive motivation that he has no need for a concept on reinforcement. Indeed, he has stated elsewhere (Logan, 1969) that reinforcement is a factor in performance and not in learning. The animal does not learn in the sense of acquiring habit strength (or losing it) because of positive reinforcement (or punishment). Rather, these response consequences increase or decrease incentive motivation, which attracts or repels the animal or makes it engage in appropriate behavior, such as approaching or withdrawing. The data Logan presents in support of this conception do not do justice to it. Hopefully there will be more and better data.

A final theoretical development comes from Estes. Estes has now joined the ranks of the incentive theorists! Appetitive behavior is based (as before) upon the sampling of stimuli to which responses have been previously conditioned, but now it is also assumed to depend upon the sampling of certain "multiplier elements". These multiplier elements are called that because they increase in weight with the severity of the animal's deprivation, and they are assumed to multiply the sampling probability of previously conditioned stimuli. These multiplier factors are not called incentive, but that is what they are, because they have all of the appropriate conceptual properties, including the possibility of innate associations with particular responses, such as consummatory responses. As though this were not enough, Estes dazzles us once again with the idea that the competition that produces suppression occurs not at the response level but at a motivational level. The emotional state produced by a noxious stimulus competes with the motivation for the baseline response by inhibiting the appropriate multiplier. To put it colloquially, the classical response-competition model of Estes and Skinner stated that the animal was too busy freezing to press the bar for food. The new Estes model tells us that the animal is too busy being frightened to be hungry. This is a new conception; I hope something comes of it.

There is a pair of chapters that have more to do with avoidance than with punishment. One by Maier, Seligman, and Solomon tells us a good deal about the noncontingent manipulation of fear and the inhibition of fear as inferred from baseline avoidance behavior, and a little about what they call helplessness, which is what happens to dogs in an aversive situation with which they cannot cope. The study of both of these new phenomena has advanced so rapidly in the last few years that the present chapter has to be seen primarily as a historical introduction. The same can be said for the chapter by Miller and Weiss, which tells us a good deal about the nature of the response requirement in avoidance learning and how to get responses other than freezing, and a little about the conditioning of autonomic responses. The latter topic presents a number of interesting theoretical issues which they touch on briefly but which I will not discuss here. I would like to make one observation, however, which is that it is one thing to show that autonomic responses

can be controlled with operant procedures, and it is quite another to show that noncontingent, respondent, or Pavlovian procedures are not conceptually valid, practically useful, and highly effective. As long as we can condition respondents (or fear, or anticipation of shock) in a trial or two independently of our animals' behavior, we are justified in doing so and in assuming that such learning is based upon separate laws of learning.

Finally, there is a pair of chapters by Mowrer and Seward that deal with the concepts of guilt and conflict, respectively, and point to some possible applications in the understanding of maladaptive human behavior.

In sum "Punishment . . ." is a fine book; there is something here for everbody, some good problems, much wisdom, and some splendid models of how to conduct a research project. There is some excellent theorizing and some fine collections of data for those who prefer to go that way.

## REFERENCES

Azrin, N. H. Some effects of two intermittent schedules of immediate and non-immediate punishment. *Journal of Psychology,* 1956, **42,** 3-21.

Campbell, B. A. and Teghtsoonian, R. Electrical and behavioral effects of different types of shock stimuli on the rat. *Journal of Comparative and Physiological Psychology,* 1958, **51,** 185-192.

Church, R. M., Wooten, C. L., and Matthews, T. J. Contingency between a response and an aversive event in the rat. *Journal of Comparative and Physiological Psychology,* 1970, **72,** 476-485.

Daly, H. B. Learning of a hurdle-jump response to escape cues paired with reduced reward or frustrative nonreward. *Journal of Experimental Psychology,* 1969, **79,** 146-157.

Leitenberg, H. Is time-out from positive reinforcement an aversive event? A review of the experimental literature. *Psychological Review,* 1965, **64,** 428-441.

Logan, F. A. *Fundamentals of learning and motivation.* Dubuque, Iowa: W. C. Brown, 1969.

Staddon, J. E. R. Temporal effects of reinforcement: a negative "frustration" effect. *Learning and Motivation,* 1970, **1,** 227-247.

Seligman, M. E. P. Chronic fear produced by unpredictable electric shock. *Journal of Comparative and Physiological Psychology,* 1968, **66,** 402-411.